

SELF-, PEER-, AND TEACHER- ASSESSMENT IN SPEAKING

Henny Rininta Awwalien

English Language Teaching Study Program

Postgraduate School, Universitas Islam Malang, Indonesia

Email: hennyrininta9@gmail.com

Abstract

This research aims to examine the correlation among self-, peer-, and teacher-assessment in speaking. It practiced a quantitative research approach, designed as correlation study. Data for this study were the score of self-, peer-, and teacher-assessment in speaking, obtained from thirty students of senior high school level. In analyzing the correlation, the researcher applied multiple regression analysis and spearman's rho correlation coefficient formula. The result of multiple correlation analysis can be stated that self- and peer-assessment have a correlation coefficient .607 with level of significance .002 means that self-assessment and peer-assessment simultaneously have an influence to teacher-assessment. The result of bivariate correlation analysis can be summarized that self-assessment and teacher-assessment have correlation coefficient .460 with level of significance .011 means that they are significantly correlated. The bivariate correlation of peer-assessment and teacher-assessment, have correlation coefficient .112 with level of significance .555 means that they are not significantly correlated. The result can be inferred that self-assessment and peer-assessment simultaneously have an influence to teacher-assessment. The relationship of peer-assessment and teacher-assessment are not significantly correlated. The significant correlation only existed between self-assessment and teacher-assessment.

Keywords: correlation, self-assessment, peer-assessment, teacher-assessment

INTRODUCTION

Assessment and testing should be defined. According to Brown & Lee (2015), a test is a system of measuring an individual's skills or knowledge within a given context, with stress on the process and measurement principle. Tests are equipments that are generally carefully planned and have recognizable scoring methods. Tests are planned administrative methods that almost take up defined periods of time during which student performance is measured.

Assessment cannot be separated from the process of teaching and learning. Test is one of a number of possible assessment forms. As Brown (2004) defined that tests are formal procedures, usually administered within strict time limits, to sample

test-taker performance in a given domain. Assessment connotes a much more boarder thought in that the majority of the time teachers teach, they also evaluate.

Referring to Brown (2004), in comparison, evaluation is an ongoing process that covers a much broader scope. Whenever a student answers a question, makes a suggestion or tries out a new word of structure, the instructor subconsciously make assessment of the student's performance.

Early in the decade of the 1990s, the conflict if idea was happened that the traditional tests could assess all individuals and all skills, a new concept then appeared which was labelled as "alternative" assessment. In line with the opinion of some expert, Brown & Lee (2015:526) explained that several alternative assessment are portfolio, journals, conferences, observations, self- and peer- assessment.

Alternatives in assessment are not the same as alternative assessment, a term that conveys the wrong message. To speak of alternative implies something that is on the periphery or "exempt from the requirements of responsible test construction" (Brown & Hudson, 1998:657). Instead, alternatives in assessment recognize that tests are one of many possible methods or design within the superordinate concept of assessment.

Sometimes such innovations are referred to as alternative assessment, if only to differentiate them from traditional formal tests. (Brown & Hudson, 1998) as cited in (Brown & Lee, 2015) explained that "alternative in assessment, which emphasizes the responsibility to apply all assessment principles to such options, and not to treat them as strange aberrations of normal assessment practices". Some alternatives in assessment are portfolios, journals, conferences, observations, and self- and peer-assessment.

Brown & Lee (2015:507) described the differences of traditional test and alternative in assessment that traditional test have characteristics such standardize exams, timed, multiple choice format, decontextualized test item, score suffice for feedback, norm-referenced score, focus on the right answer, summative, oriented to product, non-interactive performance, and foster extrinsic motivation. Whereas alternative in assessment have features such continuous long-term assessment, untimed, free-response format, contextualized tasks, formative, interactive feedback, criterion-referenced scores, open-ended, creative answers, formatives, oriented to process, interactive performances, and fosters intrinsic motivation.

Looking the transformation of school curriculum in Indonesia, the government programmed the curriculum 2013. In the field of assessment system, this curriculum implement authentic assessment. Assessment in the curriculum 2013 is referring to the regulations of department of educational and culture number 66 in the years 2013 about standard of education assessment. The standard assessment aims to: 1)

planning the students' assessment which appropriate to the achieved competence and based on the assessment principle; 2) the students' assessment is implemented professionally, open, educative, effective, efficient, and suits to the social and cultural context; and 3) the students' assessment result is reporter objectively, accountable, and informative.

Three main points in authentic assessment are attitude assessment, competence assessment, and skill assessment. Attitude assessment includes observation, self-assessment, peer assessment, journal, and interview. While competence assessment covers written test, oral test, and assignment. Skill assessment involve the performance assessment, project, and portfolio.

In this study, the writer concern in the viewpoint of self- and peer-assessment, both its implementation and contribution to students' speaking skills. Bound (1995) in Ashraf & Mahdinezhad (2015) explained that "self-assessment requires students to reflect on their own work and judge how well they have performed in relation to the assessment criteria. In other words, it provides some opportunities to be able to identify what constitutes a good piece of work".

Brown (2004) said that "peer assessment is simply one arm of a plethora of task and procedures within the domain of learner-centered and collaborative education". Falchikov (1995) in Sluijsmans et al (1998:14) defined peer-assessment as the process whereby groups or individuals rate their peers. Peer- assessment technique require students to speak then assessed by peers.

Many studies about self and peer-assessment have conducted. Orsmond et al., 2004; Ashraf & Mahdinezhad, 2015 proved that peer- and self-assessment are extremely useful in helping students increase their learning goal. On the other hand, the last study concluded that peer-assessment has more significant effect on the autonomy and speaking skills of EFL learners than self-assessment.

A study about the validity and reliability of self-assessment was conducted by Mistar (2011) investigated whether self-assessment contains build-irrelevant diversity of gender and age, and whether self-assessment relates with test score. The research found that gender and age did not create significant gaps in the validity of self-assessment.

A correlation study conducted by Zakian (2012) intended to explore assessment in English classes for the goal of contentment language learners with their marks through comparing the marks given by three sets of assessors (self-, peer-, and teacher-assessment). It found that there was a solid relationship between self-, peer-, and teacher-assessment can be valued through the formula of Pearson Product Moment Correlation. Connecting students in assessment activity creates the testing situation safe and pressure free. It should be noted for self-assessment that the

students had trouble determining their own ranks of ability. They were interested in doing SA for learning purposes, rather than testing. They couldn't just give their own presentations a mark. High correlation between teacher and student-assessment confirmations that the teacher assessment may be complemented by student assessment from a productive language skills perspective.

Another study which relevant to the self-, peer, and teacher-assessment was done by Hidayat (2013). The study focused on exploring the correlation among self-, peer-, and teacher-assessment in translation course. This research seeks the response whether self-, peer-, and teacher-assessment are correlated and to find out the confirmation of the use of self- and peer-assessment as the alternative for teacher-assessment in translation course. The result was confirmed that students think they deserve more than they were assessed. Their self-judgment proved so. However, the contributing factors causing the higher student self-assessment still calls for further research. It also confirms the assumption on students' ability in assessing their peer work properly. Unlike the self-assessment, peer-assessment seems to provide better alternative apart from teacher assessment.

One of relevant study was conducted by Edo (2017). The study purposed to describe, analyze and compare self-, peer-, and teacher-assessment in English class both quantitatively and qualitatively. The result can be emphasized that quantitatively, the result of self- and teacher-assessment show to be the most diverse, whereas self- and peer- assessment seem to be the most identical.

A study closely related to the present correlation study was a quasi-experimental study who conducted by Nida in 2017. It aimed to investigate whether the use of peer-assessment improves students' oral presentation skills and to explore students' perspectives towards the implementation of peer assessment in learning oral presentation. This study gained the data from test and questionnaire of 34 participants in Islamic Private Senior High School (MAS) Imam Syafi'i Aceh Besar. The finding confirmations that the post-test score is higher than the pre-test score. It shows that there was a significant improvement of students' skills. Moreover, questionnaire responses state that students' perspectives on the use of peer-assessment show the progressive feedback. In conclusion, the use of peer-assessment is beneficial and influential to improve students' oral presentation skills.

In accordance with the needs to solve the problem, the researcher interested to prove the relationship of self-, peer-, and teacher-assessment. This research attempt to measure the affiliation among self-, peer-, and teacher-assessment in speaking.

METHOD

This study is designed as correlation research which purposed to measure the relationship among self-, peer-, and teacher-assessment in speaking skill. According to the rules of research, this study belongs to the part of quantitative research which measure the correlation of three continuous variables; self-assessment (X1), peer-assessment (X2), and teacher assessment (Y).

This research hold in one meeting where a set of listening exercise and speaking performance were held. Thirty students of senior high school level were involved as participants. The activity was done to get the data of the correlation among self-, peer-, and teacher-assessment of students' speaking performance. Finally, the result was explained as the result of statistical calculation.

Research Instrument

In this study, assessment rubric sheets were the instrument to get the data which consist of three types; self-asesment form, peer-assessment form, and teacher-assessment form. Assessment rubric sheets are distributed to each student and completed by themselves and their peers. The teacher also complete the assessment rubric sheet on teacher-assessment form.

The self-, peer-, and teacher rating sheets are based on a journal article by Auda (2013). The scoring rubric then was modified to make the language simpler which consist of fluency, grammar, pronunciation and voice, and vocabulary. The example of scoring rubric for speaking performance is written as follow.

Table 1. Points for Speaking Performance

Criteria	4	3	2	1
Fluency	continuous speech with almost no pauses or hesitation	continuous speech with some pauses to search for adequate	frequent pauses to organize thought and/or for lack of	frequent long pauses with incomplete thought
Grammar	no or almost no errors	some minor errors that do not obscure meaning	many errors that do not obscure meaning	major errors that obscure meaning
Pronunciation & Voice	no or almost no errors	some minor errors that do not affect communication	many errors that do not affect communication	major errors that affect comprehension

	appropriate use of intonation and rhythm	some errors in using intonation and rhythm	inappropriate use of intonation and rhythm	use of mother tongue intonation and rhythm
	loud & attractive voice addressed	loud voice but not attractive	low & unattractive to whole group	quite low & unattractive voice to near colleagues
	rich use of handout vocabulary	some use of handout vocabulary	rare use of handout vocabulary	misuse of words for lack vocabulary
Vocabulary	no use of native language	rare use of native language	use of native language for about half speech	frequent use of words and phrases from the native language

To avoid misunderstanding, the scoring rubric sheet which given to the students were translated into students' first language. Before the sheets were employed by students, what each point means and how to use the scale had been clearly explained to them. To every point, the participants are asked to respond by checking the options provided to sign how good they can perform with 1 being "novice", 2 "intermediate", 3 "advanced", 4 "superior".

Data Collection

The test was held to collect the data in one meeting with the duration 3 x 45 minutes. Test refers to having students' speaking performance to certain topic given. Students were given a set of listening exercise to stimulate the speaking material. At the first step, the students received the sheets of self- and peer-assessment and they got explanation from the researcher about how to rate their peers. The main goal of this activity was to describe the assessment principles.

The students then had a listening exercise to stimulate their speaking performance. Here the students received one sheet of which contain of some questions related to what they heard from the listening recording. Participants listened to a story which is played twice. After finished with their listening section, the participants answered some questions and correct each other to their peer. The peer is one of their close friends who know well about his/her ability.

As the main view of this study, the students then continued to speaking activity. Here they delivered a speaking performance for 3 - 4 minutes individually. Students retell a story as what they heard in the listening section. Because of the limited time given, participants were divided into four groups to record their individual presentation by using a recorder tool. After all presentation has finished, self- and peer-assessment process was conducted. Students receive two sheets of scoring rubric, one was self-assessment sheet and another was peer-assessment rubric sheet.

According to the title of this research, a teacher-assessment is also needed to measure the correlation of all data. The teacher-assessment was conducted by researcher out of the class activity. Teacher assessment was done based on the students' presentation recording.

Data Analysis

The data were basically gained from the score of students' speaking performance. The correlation analysis focused on measuring the scores of self-, peer-, and teacher-assessment. To make this step easy, firstly the data are classified in the tabulation, then analyze them into the statistic program of SPSS 16. In examining the formulated hypothesis, the formula of multiple regression and spearman's rho correlation coefficient was applied. The reason of using these types of analysis are to know the multiple correlation and the bivariate correlation among the variables.

FINDING

The finding presents the results of the data analysis and the interpretation which is derived from the analysis. Data analysis presents the result of multiple regression analysis and spearman's rho correlation coefficient formula. As said before, three research problems are investigated in this research and the results are described as follows.

Multiple Correlation

To explain the multiple correlation of self- and peer-assessment to teacher-assessment, the data was analyzed by using multiple regression analysis. This technique presents four tables which are followed by the descriptions of the result.

Table 1. Descriptive Statistics among Self-, Peer-, and Teacher-Assessment

	Self- Assessment	Peer- Assessment	Teacher- Assessment
N Valid	30	30	30
Missing	0	0	0
Mean	68.83	74.53	63.77
Standard Deviation	19.700	16.141	16.519
Minimum	31	31	37
Maximum	100	100	87

The statistic table above shows that the research subject was thirty students. The subjects are completely as written in chapter 3. Mean score is 68.83 for self-assessment, 74.53 for peer-assessment, and 63.77 for teacher-assessment. In measuring self-assessment, standard of deviation is at number 19.700 with the minimum score 31 and the maximum score is 100. While peer-assessment, standard of deviation is at number 16.141 with the minimum score 31 and the maximum score is 100. In the other hand, measuring teacher-assessment show that standard of deviation which is obtained is 16.519 with the minimum score 37 and the maximum score is 87.

Table 2. Normality Test of Self-, Peer-, and Teacher-Assessment

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Self-Assessment	.170	30	.027	.930	30	.048
Peer-Assessment	.233	30	.000	.884	30	.003
Teacher-Assessment	.224	30	.000	.863	30	.001

a. Lilliefors Significance Correction

This explanation is written based on the output of table of Normality Test which aimed to find whether the data is normally distributed or not. Here the data is provided in two kinds of table; Kolmogorov-Smirnov and Shapiro-Wilk.

The Normality Test of Kolmogorov-Smirnov presents; data of self-assessment has a significance value .027, data of peer-assessment has a significance value .000, and data of teacher-assessment has a significance value .000. In this type of test, the data of variable self-assessment is normally distributed, while data of variables peer-assessment and teacher-assessment are not normally distributed.

The Test of Shapiro-Wilk presents; data of self-assessment has a significance value .048, data of peer-assessment has a significance value .003, and data of teacher-assessment has significance value .001. In this kind of test, the data of variable self-assessment is normally distributed, while data of variables peer-assessment and teacher-assessment are not normally distributed.

Table 3. Coefficient Determination

Model Summary^b				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.607 ^a	.369	.322	13.601

a. Predictors: (Constant), Self-Assessment, Peer-Assessment

b. Dependent Variable: Teacher-Assessment

The result shows that the worth of correlation coefficient (R) is .607 and the value of coefficient determination (R^2) is .369 or 36.9%. It can be decided that self-assessment and peer-assessment have contribution 36.9% to the teacher-assessment, whereas the 63.1% is explained by other factors.

Table 4. Variance Analysis (F Test)

ANOVA^b						
	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2918.480	2	1459.240	7.888	.002 ^a
	Residual	4994.887	27	184.996		
	Total	7913.367	29			

a. Predictors: (Constant), Self-Assessment, Peer-Assessment

b. Dependent Variable: Teacher-Assessment

F test purposes to define the significance of the contribution of self-assessment (X1) and peer-assessment (X2) toward teacher-assessment (Y). The result of table 3 and table 4 are written that correlation coefficient (R) is .607 with level of

significance .002 means that self-assessment and peer-assessment simultaneously have an influence to teacher-assessment.

The result of data analysis shows that F value is 7.888 with the significance value 0.002. F table value with a probability $\alpha = 0.05$ for df N1 = 2 and df N2 = 27 is 3.35. The results can be concluded that the F value (7.888) > F table (3.35), so it can be concluded that both of independent variables (self-assessment and peer-assessment) simultaneously have an influence to dependent variable (teacher-assessment).

Bivariate Correlation

To explain the bivariate correlation of self-assessment to teacher-assessment and peer-assessment to teacher-assessment, the data was analyzed by using spearman's rho correlation coefficient formula. This technique presents a table which is followed by the descriptions of the result.

Table 6. Correlation of Self-, Peer-, and Teacher-Assessment

			Self- Assessment	Peer- Assessment	Teacher- Assessment
Spearman's rho	Self-assessment	Correlation Coefficient	1.000	.123	.460*
		Sig. (2-tailed)	.	.518	.011
		N	30	30	30
	Peer-assessment	Correlation Coefficient	.123	1.000	.112
		Sig. (2-tailed)	.518	.	.555
		N	30	30	30
	Teacher-assessment	Correlation Coefficient	.460*	.112	1.000
		Sig. (2-tailed)	.011	.555	.
		N	30	30	30

*. Correlation is significant at the 0.05 level (2-tailed).

The result of bivariate correlation between self-assessment and teacher-assessment can be described that the correlation coefficient .460 with level of significance .011 means that they are significantly correlated.

As written in table 6, the result of bivariate correlation between peer-assessment and teacher-assessment can be explained that the correlation coefficient .112 with level of significance .555 means that they are not significantly correlated.

DISCUSSION

The discussion focuses again on the questions of this correlation research. The main issue is investigating the correlation among self-, peer-, and teacher-assessment in speaking. To support that, the other factors which may influence the result of this research are also explained.

The multiple correlation analysis can be defined that self- and peer-assessment have a correlation coefficient .607 with level of significance .002 means that self-assessment and peer-assessment simultaneously have an influence to teacher-assessment.

From the result of bivariate correlation analysis, can be explained that self-assessment and teacher-assessment have correlation coefficient .460 with level of significance .011 means that they are significantly correlated. The bivariate correlation of peer-assessment and teacher-assessment, have correlation coefficient .112 with level of significance .555 means that they are not significantly correlated.

From the result of multiple correlation and bivariate correlation analysis can be summarized that self-assessment and peer-assessment simultaneously have an influence to teacher-assessment. The relationship of peer-assessment and teacher-assessment are not significantly correlated. The significantly correlated only show between self-assessment and teacher-assessment.

The result of this study supports the previous study of Zakian et al (2012). The study found the high relationship between teacher- and student-assessment shows that teacher assessment can be supplemented with student assessment in the context of English skill performances. It also support the study of Mistar (2011) which gained a result that the students ' self-assessment scores reflect their true English skills. In other words, the students ' self-assessment scores are trustworthy measures of their English skills

Contrary to the result of this study, a research of Nida (2017) which examined the using of peer assessment to improve oral presentation skill, informed that the use of peer-assessment is useful and influential to improve students' oral presentation skills. It also happened in the study of Edo (2017) who conducted a study about self-, peer-, and teacher-assessment in EFL class. The result concluded that the response of

self- and teacher-assessment show to be the most diverse, whereas self- and peer-assessment seem to be the most identical.

The factors which may appear and influence the practice of self-assessment is that the students were difficult to decide the score for themselves. As Ross (2006) written in Shanti (2018) found that the force of self-assessment can be raised by training the students on how to assign their work.

In the result of peer-assessment process, the inaccurate scores may be obtained from students who assess and be assessed which caused by the relationship between students. However, the possibility has protected in the beginning by giving the students training to do the self- and peer-assessment. As Rachel et al. (2005) in Nida (2017) stated that the peer assessment is most powerful if all the students understand the criteria clearly. Besides, a concern understanding of the assessment norms can give higher validity to the output (Langan, 2005 in White, 2009).

CONCLUSION

The result of the data analysis claimed that there was a significant correlation between self- and teacher-assessment. In the first analysis which used multiple regression, the multiple correlation can be concluded that self- and peer-assessment have a correlation coefficient .607 with level of significance .002 means that self-assessment and peer-assessment simultaneously have an influence to teacher-assessment.

The result of bivariate correlation analysis, can be summarized that self-assessment and teacher-assessment have correlation coefficient .460 with level of significance .011 means that they are significantly correlated. The bivariate correlation of peer-assessment and teacher-assessment, have correlation coefficient .112 with level of significance .555 means that they are not significantly correlated.

From the result of multiple correlation and bivariate correlation analysis can be explained that self-assessment and peer-assessment simultaneously have an influence to teacher-assessment. The relationship of peer-assessment and teacher-assessment are not significantly correlated. The significant correlation only existed between self-assessment and teacher-assessment.

As the result of the study, the writer strongly agrees that by exercising students on how to do self and peer-assessment, it will improve the accuracy or the consistency of student's self and peer-assessment grades. The inaccurate result of self and peer-assessment can be the effect of students' lack information in practicing self and peer-assessment in the classroom activity. They do not comprehend how to do self- and peer-assessment, as a consequence, they overestimate or underestimate their own and friends ability.

REFERENCES

- Amalia, R. 2019. *Recast on Speaking Ability as Observer from Levels of Anxiety*. Unpublished thesis. Malang: English Graduate Program, Islamic University of Malang
- Arifin, Z. 2011. *Penelitian Pendidikan*. Bandung: Rosdakarya
- AlFallay, Ibrahim. 2004. The Role of Some Selected Psychological and Personality Trait of the Rater in the Accuracy of Self- and Peer-assessment. *English Department Journal*. Saudi Arabia: College of Arts, King Saud University
- Ashraf, Hamid & Marziyed Mahdinezhad. 2015. *The Role of Peer-assessment versus Self-assessment in Promoting Autonomy in Language Use: A Case of EFL Learners*. (Online). Volume 5, Number 2
- Auda, Zainab Jaafar. 2013. *The impact of Using Scoring Rubric in Peer Assessment on Promoting Iraqi EFL Learners' Speaking Skill in the University Level*. (Online). http://www.researchgate.net/profile/Zainab_Jaafar, accessed 14 July 2018)
- Bashir, Marriam. 2011. Factor Effecting Students' English Speaking Skills. *British Journal of Arts and Social Sciences*. (Online). Volume 2, Number 1, (http://www.bjournal.co.uk/paper/bjass_2_1/bjass_02_01_04.pdf, accessed 27 December 2017)
- Brown, H. Douglas. 2004. *Language Assessment: Principles and Classroom Practice*. New York: Pearson Education.
- Brown, H. Douglas & Lee, Heekyeong. 2015. *Teaching by Principles: An Interactive Approach to Language Pedagogy*. United States of America. Pearson Education
- Falchikov, N. 2003. *Involving Students in Assessment*. Psychology Learning and Teaching, 3(2), 102-108.
- Hidayat, M. Taufik. 2013. *Self, Peer-, and Teacher-assessment in Translation Course*. School of Postgraduate Studies Indonesia University of Education. (Online) accessed 12 May 2018.
- Kunandar. 2014. *Penilaian Autentik: Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013*. Jakarta. Rajawali Pers
- Latief, M.Adnan. 2016. *Research Methods on Language Learning: An Introduction*. Malang: UM Press
- Mistar, Junaidi. 2011. A Study of the Validity and Reliability of Self-Assessment. *Teflin Journal*, (Online), Vol 22, No 1, (<http://journal.teflin.org>, accessed 1 July 2019)
- Nida, Fastabiquil Khairat. 2017. *The Use of Peer Assessment to Improve Oral Presentation Skill*. Unpublished thesis. Banda Aceh: Faculty of Tarbiyah and Teacher Training, Ar Raniry State Islamic University
- Priyatno, Duwi. 2013. *Analisis Korelasi, Regresi dan Multivariate dengan SPSS*. Yogyakarta: Gava Media

- Santhi, Aditya Fatmala. 2018. *Intercorrelation among Self, Peer, and Teacher-Assessment of Writing Skill*. Unpublished thesis. Malang: English Graduate Program, Islamic University of Malang
- Wartim. 2018. *Speaking Assessment of Beginning Level Learners: a Case Study at Pyramid English Course Pare Kediri*. Unpublished thesis. Malang: English Graduate Program, Islamic University of Malang
- White, Eddy. (2009). *Student Perspectives of Peer Assessment for Learning in a Public Speaking Course*. Asian EFL Journal – Professional Teaching Articles. (Online), Volume 33(1), 1-36
- Yinjaroen, Pilantana & Chiramanee, Thanyapa. *Peer Assessment of Oral English Proficiency*, (Online), accessed 20 April 2018.
- Zakian, M., Moradan, A., Naghibi, S.E. 2012. *The Relationship between Self-, Peer-, and Teacher-Assessment of EFL Learners' Speaking*. Journal of Arts, Language, and Social Sciences. (Online), Volume 1, 2012, accessed 12 May 2018.